

# Estimating Correlation Energy of Diatomic Molecules and Atoms with Neural Networks

GERALDO MAGELA E SILVA, PAULO HORA ACIOLI,  
ANTONIO CARLOS PEDROZA

*Departamento de Física, Universidade de Brasília, 70 910 900 Brasília DF, Brazil*

*Received 7 November 1996; accepted 25 February 1997*

**ABSTRACT:** The electronic correlation energy of diatomic molecules and heavy atoms is estimated using a back propagation neural network approach. The supervised learning is accomplished using known exact results of the electronic correlation energy. The recall rate, that is, the performance of the net in recognizing the training set, is about 96%. The correctness of values given to the test set and prediction rate is at the 90% level. We generate tables for the electronic correlation energy of several diatomic molecules and all the neutral atoms up to radon (Rn). © 1997 by John Wiley & Sons, Inc. *J Comput Chem* 18: 1407–1414, 1997

**Keywords:** neural network; network; correlation energy

## Introduction

Neural networks have become powerful tools for researchers in recent years<sup>1</sup>. These computer models mimic the functioning of the brain for several tasks. Problems of classification, modeling, mapping, association, and dynamical processes can be handled very well, frequently giving better results than other conventional techniques. The applicability of neural networks in computa-

tional chemistry has been proved in problems of reactivity of chemical bonds, electrophilic aromatic substitution reaction, infrared spectrum–structure correlation, and in various other applications.<sup>2</sup>

Neural networks are computer programs in which the role of neurons is played by processing elements (PE) that are connected in a net and can receive and transmit information among themselves. All information received by one PE is processed with the addition of a transfer function and then transferred to other PEs. Communication between PEs is mediated by weights that work very much like the synapses between neurons. It is the modification of these weights that enables the net-

Correspondence to: G. Magela e Silva; e-mail magela@helium.fis.unb.br

work to learn and execute the tasks it is devised to do.

The most common neural networks have a multilayer architecture. The first layer is the input layer used to feed the network with the information that will model the weights in a first stage (the so-called learning process). In the second stage, when the learning process has been finished, the input layer is used to trigger the network to produce answers, referred as recalls and predictions.

The bulk of the processing is done in the layers between the input layer and the last one. Therefore, neural networks are trained with examples from where we have an output associated with a given input. After successful training the net can recall (give the correct known answers) and predict answers that have not been presented to it previously.

Accurate determination of the correlation energy of atomic and molecular systems from *ab initio* calculations requires heavy computational work, limiting the size of the systems that can be treated. Although the correlation energy is a small fraction of the total energy of the system, typically of the order of 1%, it can be of the order of chemically important energies such as binding energies and electron affinities.

In this work we use the backpropagation method of supervised learning to train a neural network to determine the correlation energy of diatomic molecules and neutral heavy atoms. In the training phase, the input set for the molecules includes the atomic numbers of the composing atoms ( $Z_1$  and  $Z_2$ ) and their bond length. For the atoms we used the Hartree–Fock energy, the atomic number, and the column on the periodic table of atoms whose correlation electronic energies are already known. Of course, other representations are possible, but this choice yields a good learning rate and a prediction rate that seems reasonable. Using the trained network we generate the correlation energies for all elements up to radon ( $Z = 86$ ). Our results are compared with the existing provisions given by other methods in the literature, as well as experimental values, when available. Our calculated correlation energy is in fairly good agreement with the experimental data and the predicted values seem very reasonable. The advantage of using the neural network is that it can calculate correlation energies of many-electron systems without prohibiting computational implications.

In the next section we make a brief review of the back propagation method and describe the network architecture used in this work. We then

discuss the important role of correlation energies in the electronic structure and the difficulty in obtaining accurate correlation energies for many-electron systems. The final section is devoted to the application of the back propagation method to model the correlation energy and tabulate values of these energies for diatomic molecules and heavy atoms. Also, we compare our estimation with other theoretical approaches and provide concluding remarks.

## Backpropagation Method and Network Architecture

Here we present a brief discussion of the method. The backpropagation of errors is a learning method; that is, a method for the correction of weights.<sup>3</sup> The great appeal of this method comes from an explicit and well-defined set of equations for the correction of the weights. The equations are applied to the correction of the weights in the last layer and successively to the previous layers up to the first. This supervised learning process utilizes pairs of input–output data. These pairs are sets of real variables, where we initially associate a known output for each given input. To begin, the weight values are set at random. The correction in this method is based on the difference between the actual response of the net and the desired output. The correction, following the gradient descent method, is computed from<sup>4</sup>:

$$\Delta W_{ji}^l = \eta \delta_j^l out_i^{l-1} + \mu \Delta W_{ji}^{l(previous)} \quad (1)$$

where  $\Delta W_{ij}^l$  represents the correction to the weight between the  $j$ th PE in the  $l$ th layer and the  $i$ th PE in the previous layer;  $out_i^{l-1}$  is the output of the  $i$ th PE on the  $l - 1$  layer;  $\eta$  and  $\mu$  are constants called the learning rate and the momentum constant, respectively. These two constants determine the rate of convergence of the learning process. In the present work we vary their values from 0.1 to 0.9 independently during the learning process to obtain the best convergence rate. The  $\delta_j^l$  values are the errors introduced by the  $j$ th PE in the  $l$ th layer. They are calculated from<sup>4</sup>:

$$\delta_j^{last} = (y_j - out_j^{last}) out_j^{last} (1 - out_j^{last}) \quad (2)$$

$$\delta_j^l = \left( \sum_{k=1}^r \delta_k^{l+1} W_{kj}^{l+1} \right) out_j^l (1 - out_j^l) \quad (3)$$

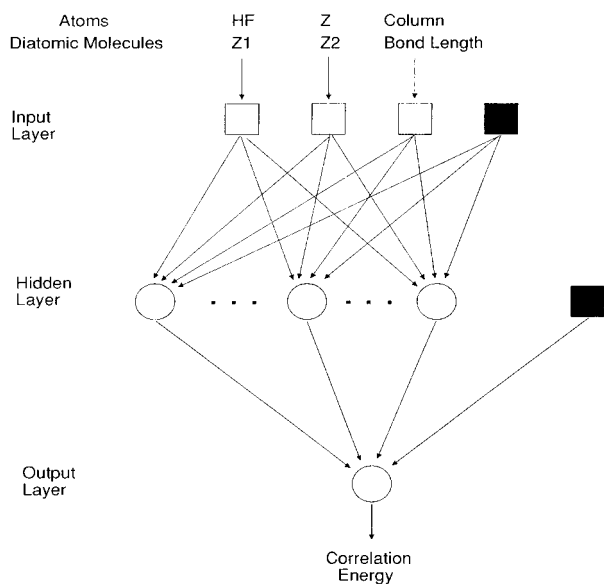
$l = 1, \dots, last - 1$

where  $y_j$  is the output set. Hence, for the correction of each weight it is necessary to use the values of three layers. We used normalized data to prevent numerical calculation problems. Upon repeatedly presenting the input and desired output sets to the network the weights are gradually corrected until the response and the desired output agree to some extent. This is accomplished by calculating the root-mean-square error between them until it becomes smaller than  $10^{-3}$  for the training set.

In our network architecture each layer of neurons is fully connected to the layer below it (Fig. 1). We considered nets with several configurations and found that the best results were given by a three-layer network. The input layer has four PEs. Three PEs identify the system. The fourth PE is a bias. The hidden layer consists of four PEs: three active PEs and a bias. The bias are introduced to formally account for the parameters of the transfer functions inherent with artificial neurons. Finally, the output layer has only one PE, giving the electronic correlation energy for the input molecule or atom.

## Correlation Energy

The correlation energy is defined as the difference between the exact nonrelativistic ground state and the Hartree–Fock energies. This small amount



**FIGURE 1.** Schematic representation of the neural network architecture used in this work.

of energy, however, is enough to cause errors in predicting the nonexistence of most negative ions or the instability for some molecules.<sup>5</sup> Also, the correlation energy is necessary for studying the electronic structures of clusters or solids.

The correlation energies obtained from very accurate experimental total energies are available only to light atoms<sup>6</sup> and small molecules.<sup>7–9</sup> No reliable experimental data are available for many-electron systems.

To obtain measurement of theoretically accurate correlation energy, one needs to perform heavy computational calculations such as that done in configuration interaction or many-body perturbation methods<sup>5</sup> with relativistic considerations. These methods, however, are limited to systems with no more than two dozen or so electrons. Almost equal accuracy is obtained using the density functional theory within its nonlocal approximations.<sup>10</sup> This approach needs less computational effort and, thus, results for a few heavy atoms have been published<sup>10</sup>. Using a semiempirical model Chakravorty and Clementi<sup>11</sup> made very accurate estimation of the correlation energy for many atoms. In their study, however, they did not go beyond the Xe atom.

In the treatment of diatomic molecules Carrol et al. used local (including self-correlation corrections) and nonlocal correlation energy functionals. They provided an extensive table of diatomic molecule correlation energies<sup>7</sup>. Savin et al. also considered the nonlocal approximation to generate the correlation energy of several molecules<sup>8</sup>. The use of many-body perturbation and coupled cluster methods has also been taken into account in the determination of the correlation energy of diatomic molecules.<sup>9</sup>

## Discussion and Summary

In this work we used a very flexible method to estimate the correlation energies of diatomic molecules and heavy atoms. During the learning process we taught the neural network the exact known correlation energies and bond lengths.<sup>6–9</sup> The variables in the input for the diatomic molecules are the atomic numbers of the atoms ( $Z_1$  and  $Z_2$ ) and their bond length. The network was trained using a set of 31 molecules and the prediction rate was determined using a set of 6 test molecules. We then used the network to predict

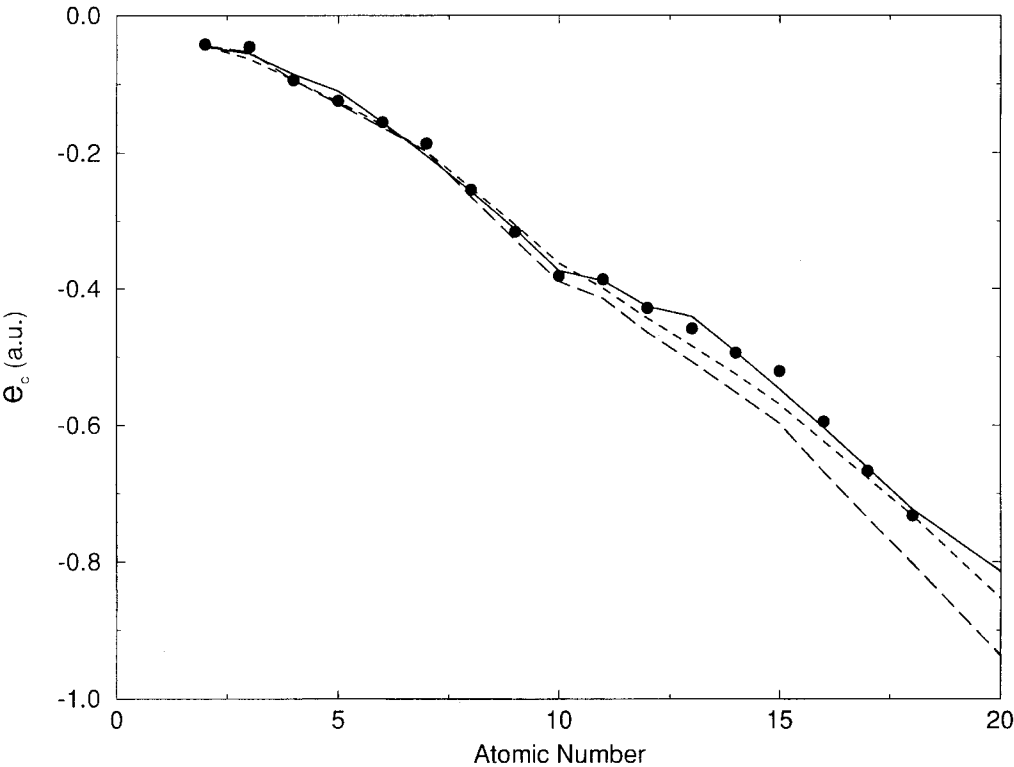
the correlation energies of 28 molecules. For the atomic case we used the atomic number and the column of the atom in the periodic table and its Hartree–Fock energy.<sup>12</sup>

One remark should be made, though—there can be a deficiency in this representation in the estimation of correlation energies of the lanthanides, as we chose to represent all of them as column 3 in the periodic table. They are distinguished by their atomic numbers and HF energies. However, Figure 3 shows that the estimation seems reasonable.

In Table I we present the results of the learning process. The difference between the energies predicted by the network and the experimental values<sup>6</sup> (taken as exact values) is about 4%. Overall, results are better than those estimates using density functional theory (DFT) in the local spin density (LSDA)<sup>13</sup> and generalized gradient (GGA)<sup>13</sup> approximations. We can better visualize the difference in Figure 2 where we plot the LSDA (dashed lines), GGA (long dashed lines), exact energies (full circles), and the energies learned by the neu-

**TABLE I.** Correlation Energies (in a.u.) for Atoms Helium through Argon (Comparison of LSDA, GGA, Neural Network [NN], and Exact Values).

Z	LSDA (Ref. 12)	GGA (Ref. 12)	NN	Exact (Ref. 5)
2	−0.045	−0.044	−0.045	−0.042
3	−0.063	−0.053	−0.055	−0.045
4	−0.094	−0.094	−0.086	−0.094
5	−0.125	−0.128	−0.110	−0.124
6	−0.159	−0.163	−0.155	−0.155
7	−0.198	−0.198	−0.204	−0.186
8	−0.251	−0.264	−0.256	−0.254
9	−0.305	−0.328	−0.312	−0.316
10	−0.362	−0.389	−0.373	−0.381
11	−0.399	−0.414	−0.388	−0.386
12	−0.443	−0.464	−0.426	−0.428
13	−0.484	−0.507	−0.441	−0.459
14	−0.525	−0.552	−0.493	−0.494
15	−0.569	−0.597	−0.547	−0.521
16	−0.623	−0.668	−0.603	−0.595
17	−0.677	−0.736	−0.662	−0.667
18	−0.732	−0.802	−0.723	−0.732



**FIGURE 2.** Correlation energies (in a.u.) as a function of atomic number (Z), learning stage. (a) Exact values (full circles); (b) Predictions of the neural network (full line); (c) GGA results (dashed lines); and (d) LSDA results (long dashed lines).

**TABLE II.**  
**Estimates of Correlation Energies (in a.u.) for Neutral Atoms from Helium through Argon Using a Backpropagation Neural Network.**

Z	$e_c$	Z	$e_c$	Z	$e_c$	Z	$e_c$	Z	$e_c$
2	-0.045	19	-0.768	36	-1.581	53	-2.390	70	-2.918
3	-0.055	20	-0.813	37	-1.675	54	-2.416	71	-2.941
4	-0.086	21	-0.858	38	-1.738	55	-2.425	72	-2.990
5	-0.110	22	-0.903	39	-1.799	56	-2.490	73	-3.034
6	-0.155	23	-0.948	40	-1.857	57	-2.550	74	-3.075
7	-0.204	24	-0.994	41	-1.913	58	-2.583	75	-3.112
8	-0.256	25	-1.041	42	-1.966	59	-2.615	76	-3.146
9	-0.312	26	-1.088	43	-2.017	60	-2.647	77	-3.176
10	-0.373	27	-1.135	44	-2.065	61	-2.677	78	-3.203
11	-0.388	28	-1.183	45	-2.110	62	-2.707	79	-3.228
12	-0.426	29	-1.232	46	-2.153	63	-2.736	80	-3.249
13	-0.441	30	-1.281	47	-2.194	64	-2.764	81	-3.268
14	-0.493	31	-1.331	48	-2.232	65	-2.791	82	-3.284
15	-0.547	32	-1.381	49	-2.268	66	-2.818	83	-3.298
16	-0.603	33	-1.431	50	-2.301	67	-2.844	84	-3.310
17	-0.662	34	-1.481	51	-2.333	68	-2.869	85	-3.319
18	-0.723	35	-1.531	52	-2.362	69	-2.894	86	-3.327

ral network (full line). Of course, the DFT has a more physical justification, but the approach in this work can be equally justified as a mapping of known results that model the neural network and its *a posteriori* use to estimate the unknown correlation energies. In Table II we tabulate all correlation energies estimated in this work.

In Figure 3 we plot the correlation energies predicted by our network for all atoms up to radon ( $Z = 86$ ). One can see that the correlation energies (absolute values) increase with the atomic number, as expected. One can also note that there is a certain periodicity in the predictions, as may be expected, due to the periodic nature of the elements. The correlation energies predicted by the neural network are somewhat smaller than those predicted by LSDA and GGA.<sup>13</sup> Nevertheless, it is known that LSDA overestimates the atomic correlation energies and the GGA should correct this trend. So the neural network prediction is also coherent in this respect. The very elaborate result of Kelly and Ron<sup>14</sup> for the iron atom is also displayed in Figure 3, and is in good agreement with our results. Of course we would need additional, exact data to have a better idea of how good our estimations are or how to better train our network. Therefore, a more comprehensive list of exact correlation energies could be very useful in improv-

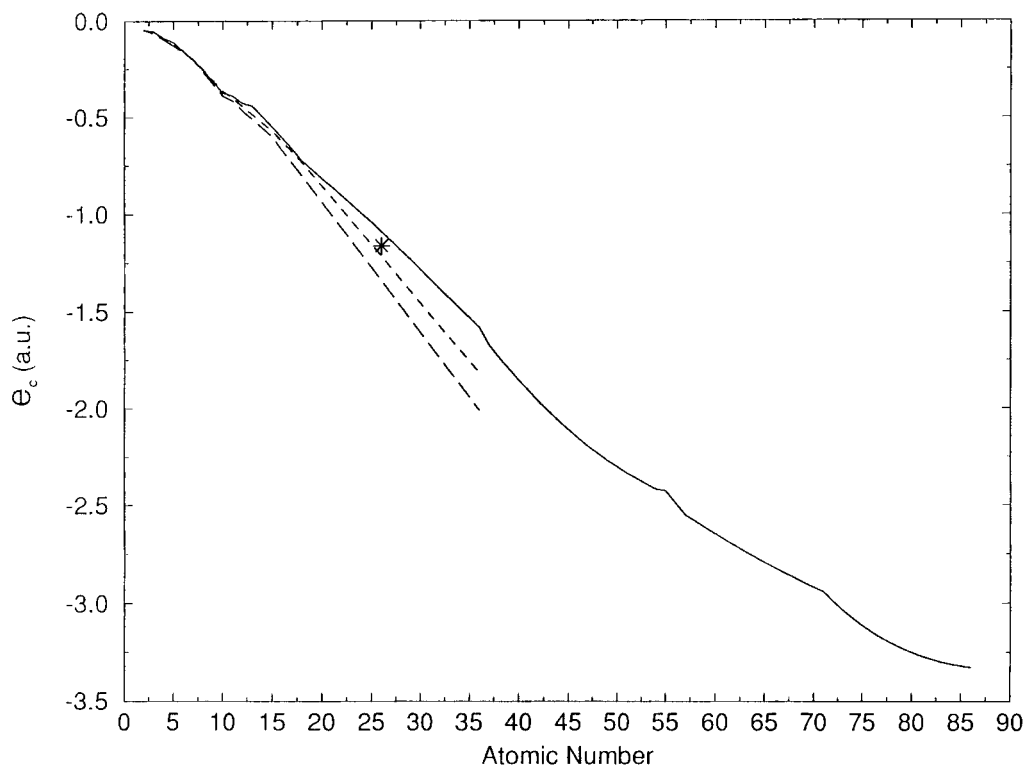
ing our current neural network. In particular we are interested in exact correlation energies for the transition metals.

It should be noted that, for high- $Z$  atoms, relativistic effects will have a profound effect on electronic correlation. Also, some neural networks have demonstrated poor reliability when the range of a given input pattern of the test set exceeds the corresponding range of the training set.<sup>15</sup>

Figure 4 presents the recognition and prediction of our learning and test sets (Table III). The overwhelming concordance of the neural network and the experimental results has stimulated us to make predictions of several unknown correlation energies of diatomic molecules. These predictions are shown in Figure 5 and Table IV. New experimental values for the molecules shown in Figure 5 are expected to be in good accordance with the present results.

It should be noted that the use of network predictions to calculate third quantities, such as the binding energy in the diatomics, could lead to rather large deviations due to error propagation.

In conclusion, we have succeeded in generating estimates of the correlation energies of several diatomic molecules and neutral atoms up to radon using a backpropagation neural network. Because there are no exact correlation energies for many-

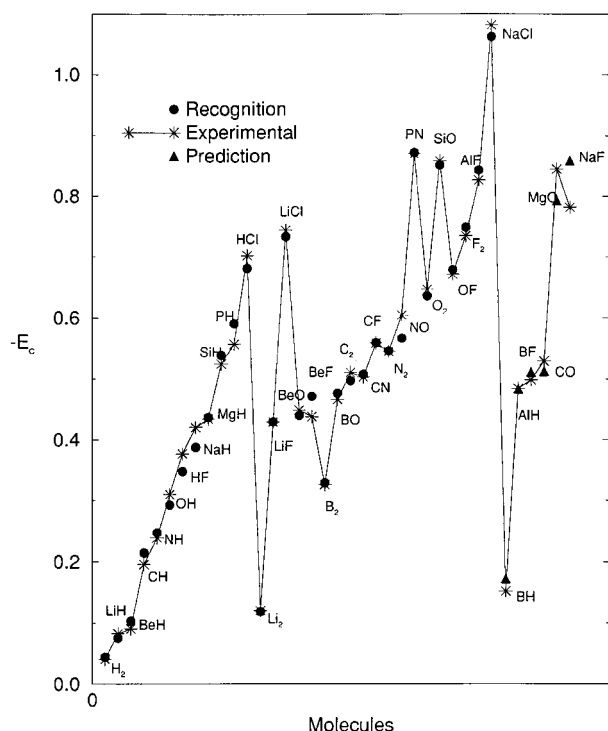


**FIGURE 3.** Correlation energies (in a.u.) as a function of atomic number ( $Z$ ), prediction stage. (a) Predictions of the neural network (full line); (b) GGA results (dashed lines); (c) LSDA results (long dashed lines); and (d) Kelly and Ron estimate for the iron atom (star).

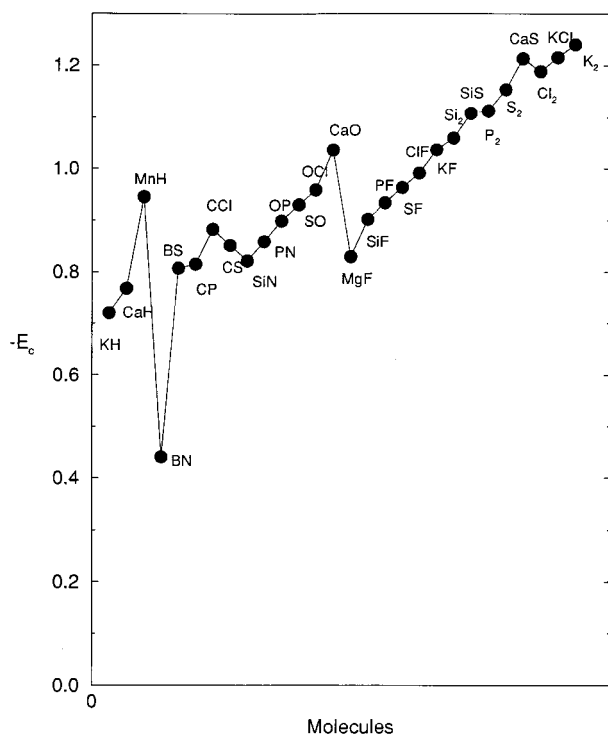
**TABLE III.**  
**Estimates of Correlation Energies of Diatomic Molecules.<sup>a</sup>**

Molecule	$e_c$ (neural network)	$e_c$ (exact)	Molecule	$e_c$ (neural network)	$e_c$ (exact)
H <sub>2</sub>	-0.044	-0.041	C <sub>2</sub>	-0.497	-0.510
LiH	-0.076	-0.083	CN	-0.508	-0.503
BeH	-0.103	-0.091	CF	-0.560	-0.559
CH	-0.214	-0.196	N <sub>2</sub>	-0.546	-0.545
NH	-0.247	-0.239	NO	-0.567	-0.605
OH	-0.293	-0.310	PN	-0.872	-0.870
HF	-0.348	-0.377	O <sub>2</sub>	-0.636	-0.647
NaH	-0.388	-0.420	SiO	-0.851	-0.858
MgH	-0.437	-0.434	OF	-0.679	-0.672
SiH	-0.539	-0.524	F <sub>2</sub>	-0.749	-0.735
PH	-0.591	-0.557	AlF	-0.843	-0.826
HCl	-0.681	-0.702	NaCl	-1.063	-1.082
Li <sub>2</sub>	-0.119	-0.120	BH	-0.172	-0.152
LiF	-0.430	-0.429	AlH	-0.483	-0.484
LiCl	-0.733	-0.744	BF	-0.511	-0.498
BeO	-0.440	-0.449	CO	-0.512	-0.529
BeF	-0.471	-0.438	MgO	-0.793	-0.844
B <sub>2</sub>	-0.329	-0.326	NaF	-0.858	-0.782
BO	-0.476	-0.466			

<sup>a</sup>The first 31 molecules (H<sub>2</sub> to NaCl) were used to train the neural network. The last six (BH to NaF) were used to determine the prediction rate of the network.



**FIGURE 4.** Correlation energies of diatomic molecules. Experimental values: stars; network recognition: full circles; predictions: full triangles.



**FIGURE 5.** Network prediction of the correlation energy of diatomic molecules.

**TABLE IV.**  
Estimate of Correlation Energies (in a.u.) for Diatomic Molecules.

Molecule	$e_c$	Molecule	$e_c$
KH	-0.720	MgF	-0.829
CaH	-0.768	SiF	-0.902
MnH	-0.945	PF	-0.934
BN	-0.440	SF	-0.964
BS	-0.806	ClF	-0.991
CP	-0.814	KF	-1.037
CCl	-0.882	Si <sub>2</sub>	-1.060
CS	-0.851	SiS	-1.108
SiN	-0.821	P <sub>2</sub>	-1.112
PN	-0.858	S <sub>2</sub>	-1.153
OP	-0.898	CaS	-1.213
SO	-0.930	Cl <sub>2</sub>	-1.189
OCI	-0.958	CKI	-1.216
CaO	-1.036	K <sub>2</sub>	-1.241

electron systems we compare our estimates with correlation energies obtained in the framework of the DFT in the LSDA and GGA approximations. Our energies are smaller (in absolute value) than those obtained in those approximations, a trend followed by the known exact results.

Extensions of this work to provide estimates of dipole moments of diatomic molecules, corrections to HF, MP2 homolytic bond dissociation energies, and some of the parameters involved in parameterizing the G2 method are currently being undertaken and the results will be presented in a forthcoming article.

## References

1. T. L. H. Watkin, A. Rau, and M. Biehl, *Rev. Mod. Phys.*, **65**, 499 (1993).
2. J. Gasteiger and J. Zupan, *Angew. Chem. Int. Ed. Engl.*, **32**, 503 (1993).
3. D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Parallel Distributed Processing: Explorations in the Microstructures of Cognition*, Vol. I, D. E. Rumelhart and J. L. McClelland, Eds., MIT Press, Cambridge, MA, p. 318.
4. J. Zupan and J. Gasteiger, *Neural Networks for Chemists*, VCH, Weinheim, 1993, pp. 122, 123, and 147.
5. S. Wilson, Ed., *Methods in Computational Physics*, Vol. I, Plenum, New York, 1987.
6. A. Veillard and E. Clementi, *J. Chem. Phys.*, **49**, 2415 (1968).
7. M. T. Carroll, R. F. W. Bader, and S. H. Vosko, *J. Phys. B Atom. Mol. Phys.*, **20**, 3599 (1987).

8. A. Savin, U. Wedig, H. Preuss, and H. Stoll, *Phys. Rev. Lett.* **53**, 2087 (1984).
9. L. Adamowicz, R. J. Barlett, and E. A. McCullough, Jr., *Phys. Rev. Lett.* **54**, 426 (1985).
10. Y.-M. Juan and E. Kaxiras, *Phys. Rev. B*, **48**, 14944 (1993).
11. S. J. Chakravorty and E. Clementi, *Phys. Rev. A*, **39**, 2290 (1989).
12. C. F. Bunge, J. A. Barrientos, and A. V. Bunge, *Atom. Data Nucl. Data Tables*, **53** 113 (1993).
13. O. V. Gritsenko, N. A. Cordero, A. Rubio, L. C. Balbás, and J. A. Alonso, *Phys. Rev. A*, **48**, 4197 (1993).
14. H. Kelly and A. Ron, *Phys. Rev. A* **4**, 11 (1971).
15. G. Rauhut, J. W. Boughton, and P. Pulay, *J. Chem. Phys.* **103**, 5662 (1995).